AFOSR-TR-97

0264

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		FINAL 01 Jun 93 To 31 Dec 96

4. TITLE AND SUBTITLE

DEMODULATION PROCESSES IN AUDITORY PERCEPTION

5. FUNDING NUMBERS

F49620-93-1-0299
61102F
2313/AS

6. AUTHOR(S)

Dr Lawrence Feth

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Speech and hearing Science
Ohio State University
110 Pressey Hall
Columbus OH 43210

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFOSR/NL
110 Duncan Avenue Room B115
Bolling AFB DC 20332-8050

Maj William P. Roach

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

19970616 087

12a. DISTRIBUTION AVAILABILITY STATEMENT

Approved for public release;
distribution unlimited.

DTIC QUALITY INSPECTED 3

13. ABSTRACT (Maximum 200 words)

The long range goal of this project was the understanding of human auditory processing of information conveyed by complex, time-varying signals such as speech, music or important environmental sounds. Our work was guided by the assumption that human auditory communication is a "modulation - demodulation" process. That is, we assume that sound sources produce a complex stream of sound pressure waves with information encoded as variations (modulations) of the signal amplitude and frequency. The listener's task then is one of demodulation. Much of past psychoacoustics work has been based in what we characterize as "spectrum picture processing." Complex sounds are Fourier analyzed to produce an amplitude-by-frequency "picture" and the perception process is modeled as if the listener were analyzing that spectral picture. This approach leads to studies such as "profile analysis" and to the power-spectrum model of masking. Our approach leads us to investigate time-varying, complex sounds. We refer to them as dynamic signals and we have developed auditory singnal processing models to help guide our experimental work.

14. SUBJECT TERMS	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
(U)	(U)	(U)	(UL)

AFOSR--93-0299

DEMODULATION PROCESSES IN AUDITORY PERCEPTION

LAWRENCE L. FETH, Ph.D.
SPEECH AND HEARING SCIENCE
OHIO STATE UNIVERSITY
110 PRESSEY HALL
COLUMBUS, OHIO 43210

1 JUNE 1997

FINAL REPORT: 1 JUNE 1993 - 31 DECEMBER 1996

PREPARED FOR:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
BOLLING AIR FORCE BASE
WASHINGTON, DC 20332

FINAL TECHNICAL REPORT

1 JUNE 1993 - 31 DECEMBER 1996

Introduction

The long range goal of this project was the understanding of human auditory processing of information conveyed by complex, time-varying signals such as speech, music or important environmental sounds. Our work was guided by the assumption that human auditory communication is a “modulation - demodulation” process. That Is, we assume that sound sources produce a complex stream of sound pressure waves with information encoded as variations (modulations) of the signal amplitude and frequency. The listener's task then is one of demodulation. Much of past psychoacoustics work has been based in what we characterize as “spectrum picture processing.” Complex sounds are Fourier analyzed to produce an amplitude-by-frequency “picture” and the perception process is modeled as if the listener were analyzing that spectral picture. This approach leads to studies such as “profile analysis” and to the power-spectrum model of masking. Our approach leads us to investigate time-varying, complex sounds. We refer to them as dynamic signals and we have developed auditory signal processing models to help guide our experimental work.

This final report covers the final three years of a ten-year project funded by AFOSR. It began while the PI was employed at the University of Kansas and continued with his move to Ohio State. Most of the significant results have been obtained at Ohio State due to the close collaboration of the PI with Ashok Krishnamurthy in Electrical Engineering. Graduate students in Speech and Hearing Science and in Electrical Engineering have been funded on the project. In turn, the students have adopted many of the sub-projects for their thesis or dissertation work. While this has been beneficial to both the faculty advisors and the graduate students, it has led to an unacceptable lag in getting research results into referred publications. The PI is currently attempting to remedy this situation by writing up some of student projects from several years ago. He does not intend to seek additional outside funding from AFOSR or other federal sources until a major part of this backlog is in press.

Sub-projects will be reviewed in the order presented in the three annual reports (1994, -95 & -96). Although the order of studies differs from the original proposal, we will again review progress under those headings to enable a comparison with that document. Our work on multi-channel IWAIF models has been influenced by the now classical work of Chistovitch on “the perceptual formant” in vowel perception. Our preliminary model did a reasonable job of predicting listener performance in the “center-of-gravity” experiments she reported twenty years ago. We were fortunate to have Jaan Ross (a student of Chistovitch) , visiting our lab in November and December of 1995. He helped us to extend the Chistovitch work from steady vowels to dynamic sounds.

We also applied some of our resources to the study of signal processing by people who have been fitted with a cochlear implant. These special cases allow us to test our notions of temporal processing in a sensory system somewhat different from the normal hearing human.

List of research objectives and cumulative progress

A. Single-transition signals - single channel model

1. Frequency Modulated Tones

a. Roving frequency: GLIDE–STEP Discrimination (see paragraph below)

b. Sinusoidal vs. Linear Trajectory

Our results were reported at the June 94 meeting of the Acoustical Society and at the 10th International Symposium on Hearing at Irsee, Bavaria. A manuscript submitted to the Journal of the Acoustical Society of America is still undergoing revisions to satisfy the comments of the reviewers. The Irsee proceedings are in print.

c. Slope Discrimination

This area was the topic of the doctoral dissertation of Chien yeh Hsu. His dissertation was completed in the summer quarter 1993, and he reports that a manuscript is in progress. From July 1993 until summer 1995, he held a post-doctoral position at the University of Illinois. He has since returned to Taiwan. A manuscript based on Hsu's work is still planned for submission to JASA, but this is one that the PI will have to write since Hsu is unavailable to work on it.

2. Moving Filter

a. Variation on the GLIDE-STEP Discrimination task

We have by-passed these proposed follow up versions of the original design because we decided that most of the questions raised could be answered by incorporating the variations into sinusoidal plus linear FM designs.

b. Sinusoidal vs. Linear Trajectory

These experiments using the "moving filter" signals were not conducted in favor of work on the "center-of-gravity" studies.

c. Slope Discrimination

The discrimination of the slope of the linear trajectory for a single resonator filter was incorporated into the dissertation of Hsu (see 1.c above). Results were reported at the 1994 meeting of the ARO.

3. Single Formants from "Real Speech"

These experiments were not conducted in favor of work on the "center-of-gravity" studies.

B. Multi-formant signals - multi-channel model

Work on the multi-channel IWAIF model was incorporated into the master's thesis of M. Mokheimer, who applied it to the detection of mixed modulation by human listeners. A presentation based on the thesis is to be given in Cairo in Dec. 1994. Development of the model has continued in the dissertation work of Jayanth Anantharaman which is expected to be finished by December 1997.

1. Moving filter multiple-formant signals. (see paragraph 2.a above)

2. Multi-formant signals extracted from real speech (see paragraph 3 above)

C. Incorporation of Envelope Cues

We have conducted a series of experiments suggested by the work of Versfeld and Houtsma, and presented preliminary results at the June 94 meeting of the Acoustical Society. As is often the case, the experiments have raised more questions than they answered. Despite our best efforts, we have failed to reconcile the differences reported by Versfeld and Houtsma and our earlier work on IWAIF predictions using simple two-component tones. A joint publication with Dai, Nguyen, Kidd and Green (see below) represents our latest efforts with respect to studies of simple IWAIF predictions using tonal signals. Also, A manuscript titled "Synthesis of Common-Envelope Signal Pairs" was revised and resubmitted to JASA.

D. Spectral Center-of-Gravity effects

In winter quarter 1996, Julie Lester began an undergraduate honors thesis to replicate and extend the perceptual formant, or spectral center-of-gravity work reported by Chistovitch in 1979. That project was continued through the final year of this project by Qiang Xu a graduate student in Speech and Hearing Science. Presentations of preliminary results were given at the Fall 1996 and the Spring 1997 meetings of the Acoustical Society. Xu completed a masters thesis that extended the perceptual linear prediction model of Hermansky (1990) to the spectral integration across formants in synthetic vowel-like sounds. By matching a single formant, variable frequency signal to a reference having two formant peaks, listeners indicate the frequency locus of the equivalent perceptual formant.

Our preliminary work replaced the simple matching task with a modified double-staircase procedure as described by Jesteadt (1980) for subjective judgments. Using that procedure, Xu has studied the effects of signal excitation type (periodic pulse train versus wideband noise), formant bandwidth, and frequency separation of the formants in the reference signal. Neither excitation type nor formant bandwidth appear to have significant effects on listener performance. Formant separation leads to what Chistovitch called the critical separation effect. She reported that when formants we more than 3.5 Bark apart, spectral integration failed. Xu's results confirm that notion and the modeling appears to account for the effect. Individual differences between Chistovitch's two subjects and among our three listeners are evident, but were not accounted for by the modeling. (Of course, modeling of individual differences was not a goal of the model).

E. Other work - cochlear implants

Ina Bicknell chose to use the STEP-GLIDE signals used in our early work to assess the temporal acuity of listeners who are implanted with Cochlear Corp. Nucleus 22 devices. This was the main focus of Bicknell's dissertation work which was designed to understand why implant users still have difficulty with many speech sounds. The dissertation work was completed in June 1996 and after a brief post-doctoral stay at the University of Michigan, Bicknell is working on preparation of manuscripts from the dissertation for submission to the appropriate journals. One of the manuscripts will focus on the STEP-GLIDE discrimination results. Briefly, implant wearers have great difficulty with the task, presumably because the design of the speech processing device that drives the implanted electrodes does not take into consideration the dynamic nature of speech formant transitions.

Gerald Lynch completed a master's thesis in Electrical Engineering in which he modified the Patterson AIP model to reflect the signal processing of the cochlear implant device. Signals such as those used by Bicknell were used to test the modeling. We expect that a manuscript with Bicknell's results may include the modeling results from Lynch's work.

Participating Professionals

Lawrence L. Feth, Ph.D.	Principal Investigator
Ashok K. Krishnamurthy, Ph.D.	Co-Investigator
**Jayanth N. Anantharaman, M.S.	Grad. Research Assoc.
Tao Zhang, M.S.	Grad. Research Assoc.
*Chien-yeh Hsu M.S.	Grad. Research Assoc.
*Mohamed Mokheimer	Grad. Research Assoc.
Gerald P. Lynch, B.S.	Grad. Research Assoc.
*Ina R. Bicknell M.S.	Grad. Research Assoc.
*Julie L. Lester	Undergraduate Research Assoc.
*Qiang Xu	Grad. Research Assoc.

(*no stipend cost to this project [**after June 1996])

Publications

Detection of frequency modulation in steady and gliding tones. T. Zhang, L. L. Feth and A. K. Krishnamurthy, Manuscript still under revision for J. Acoust. Soc. Am. (1997).

Synthesis of common-envelope signal pairs. J. A. Anantharaman, A. K. Krishnamurthy, and L. L. Feth Manuscript re-submitted to J. Acoust. Soc. Am. (1996).

Phase Independence of Pitch Produced by Narrowband Sounds. H. Dai, Q. T. Nguyen, G. Kidd, L. L. Feth and D. M. Green. J. Acoust. Soc. Amer. 100, 2349-2351 (1996).

Intensity-weighted average of instantaneous frequency as a model for frequency discrimination. J. N. Anantharaman, A. K. Krishnamurthy and L. L. Feth, J. Acoust. Soc Amer., Vol. 94, 723-729, (1993).

Complex Sound Discrimination and a Correlated Channel Model. T. Zhang Unpublished Ph. D. dissertation, The Ohio State University. (1995).

Detection of Frequency Modulation by Listeners with Cochlear Implants. I. R. Bicknell, Doctoral Dissertation, Ohio State University. (1996).

Modeling Cochlear Implant Processing Schemes. G. P. Lynch, Masters Thesis (E.E.) Ohio State University. (1996).

Spectrographic Analysis of Sound vs. Auditory Perception. J. L. Lester, Senior Honors Thesis, Ohio State University, (1996).

A Signal Processing Model for Spectral Integration of Synthetic Vowel-like Sounds. Qiang Xu, Masters Thesis (SHS) Ohio State University, (1997)

Detection of Combinations of frequency modulation: An application of the IWAIF model, L. L. Feth, A. K. Krishnamurthy and T. Zhang in Advances in Hearing Research, G. A. Manley, G. M. Klump, C. Köppl, H. Fastl and H. Oeckinghaus, editors. World Scientific, Singapore, pages 415-424, (1995).

A Multi-channel Intensity-Weighted Average of Instantaneous Frequency Model. M. Mokheimer, J. N. Anantharaman, A. K. Krishnamurthy and L. L. Feth First International Conference on Electronics, Circuits and Systems, Cairo, Egypt (Dec. 1994).

Interactions/Transitions

a) Presentations at meetings

Discrimination of broadband, multi-component, common-envelope signals. J. N. Anantharaman, A. K. Krishnamurthy and L. L. Feth, [Abstract: J. Acoust. Soc Amer. 93, p2387, (1993)].

Short-term IWAIF model for frequency discrimination. A. K. Krishnamurthy and L. L. Feth, [Abstract: J. Acoust. Soc Amer. 93, p2387, (1993)].

A two-dimensional, independent channels model for complex sound discrimination. T. Zhang, C. Hsu, L. L. Feth and A. K. Krishnamurthy, [Abstract: Seventeenth Midwinter Research Meeting of the Association for Research in Otolaryngology, p 69, (1994)].

An IWAIF model for the detection of mixed modulation. M. Mokheimer, J. N. Anantharaman, L. L. Feth and A. K. Krishnamurthy, [Abstract: Seventeenth

Midwinter Research Meeting of the Association for Research in Otolaryngology, p71, (1994)].

Discriminability of three-tone, common envelope signals. J. N. Anantharaman, L. L. Feth and A. K. Krishnamurthy, [Abstract: J. Acoust. Soc Amer. 95, p2964, (1994)].

Detection of frequency modulation in steady and gliding tones. T. Zhang, L. L. Feth and A. K. Krishnamurthy, [Abstract: J. Acoust. Soc Amer. 95, p2965, (1994)].

A correlated channels model for complex sound discrimination, T. Zhang, L. L. Feth, and A. K. Krishnamurthy, [Abstract:18th Midwinter meeting of the Association for Research in Otolaryngology (1995)].

IWAIF model predictions of a profile analysis task. M. A. Ericson, L. L. Feth and A. K. Krishnamurthy [Abstract: J. Acoust. Soc. Am. 97, pg.3272 (1995)].

On the perceptual asymmetry in frequency modulation discrimination, T. Zhang, L. L. Feth and A. K. Krishnamurthy [Abstract: J. Acoust. Soc. Am. 97, pg.3273 (1995)].

Synthesis of common-envelope signals. J. A. Anantharaman, A. K. Krishnamurthy and L. L. Feth [Abstract: J. Acoust. Soc. Am. 97, pg.3274 (1995)].

Applications of the IWAIF model to auditory processing of dynamic signals. L. L. Feth, A. K. Krishnamurthy, T. Zhang and I. Bicknell. Invited presentation at the Conference on Implantable Auditory Prostheses, Asilomar, CA August, 1995.

Intensity Weighted Average of Instantaneous Frequency Computations Using Time-Frequency Representations. J. N. Anantharaman, A. K. Krishnamurthy and L. L. Feth, [Abstract: J. Acoust. Soc. Am 99, pg.2564 (1996)].

Detection of Frequency-Modulated Signals by Cochlear-Implant Users. I. R. Bicknell and L. L. Feth [Abstract: J. Acoust. Soc. Am. 99, pg.2564 (1996)].

Testing the 'center-of-gravity' effect for vowel-like complex sounds. L. L. Feth, J. L. Lester, Q. Xu and J. Ross [Abstract: J. Acoust. Soc. Am. 100, pg.2626 (1996)].

Patents and Inventions

No patentable inventions have resulted from this research.

General Statements